



PCT

#10

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/10/009,030A

TIME: 14:33:45

Input Set : A:\Bml41018.app

Output Set: N:\CRF4\02112003\J009030A.raw

3 <110> APPLICANT: ARAND, MICHAEL
 4 ARCHELAS, ALAIN ROBERT
 5 BARATI, JACQUES
 6 FURSTOSS, ROLAND
 8 <120> TITLE OF INVENTION: EPOXIDE HYDROLASES OF ASPERGILLUS ORIGIN
 10 <130> FILE REFERENCE: bml-410.018
 12 <140> CURRENT APPLICATION NUMBER: 10/009,030A
 C--> 13 <141> CURRENT FILING DATE: 2002-12-20
 15 <150> PRIOR APPLICATION NUMBER: PCT/FR00/01217
 16 <151> PRIOR FILING DATE: 2000-05-05
 18 <150> PRIOR APPLICATION NUMBER: FR 99/05711
 19 <151> PRIOR FILING DATE: 1999-05-05
 21 <160> NUMBER OF SEQ ID NOS: 8
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 1197
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Aspergillus niger
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (1)..(1194)
 34 <400> SEQUENCE: 1
 35 atg tcc gct ccg ttc gcc aag ttt ccc tcg tcg gcg agc att tcg cct 48
 36 Met Ser Ala Pro Phe Ala Lys Phe Pro Ser Ser Ala Ser Ile Ser Pro
 37 1 5 10 15
 39 aat cct ttc acg gtc tct atc ccg gat gaa cag ttg gat gac ttg aaa 96
 40 Asn Pro Phe Thr Val Ser Ile Pro Asp Glu Gln Leu Asp Asp Leu Lys
 41 20 25 30
 43 acc ctc gtc cga ctg tcc aag att gct cct ccc acc tat gag agc ctg 144
 44 Thr Leu Val Arg Leu Ser Lys Ile Ala Pro Pro Thr Tyr Glu Ser Leu
 45 35 40 45
 47 caa gcg gat ggc cgg ttt ggc atc act tct gaa tgg ctg aca act atg 192
 48 Gln Ala Asp Gly Arg Phe Gly Ile Thr Ser Glu Trp Leu Thr Thr Met
 49 50 55 60
 51 cgg gag aag aaa tgg ctc tcg gag ttt gac tgg cga cca ttt gaa gct cga 240
 52 Arg Glu Lys Trp Leu Ser Glu Phe Asp Trp Arg Pro Phe Glu Ala Arg
 53 65 70 75 80
 55 ctg aac tct ttc cct cag ttt act aca gag atc gag ggt ctc acg att 288
 56 ~~Leu Asn Ser Phe Pro Gln Phe Thr Thr Glu Ile Glu Gly Leu Thr Ile~~
 57 85 90 95
 59 cac ttt gct gct ctc ttc tcc gag agg gag gat gct gtg cct atc gca 336
 60 His Phe Ala Ala Leu Phe Ser Glu Arg Glu Asp Ala Val Pro Ile Ala
 61 100 105 110

ENTERED

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/10/009,030A

TIME: 14:33:45

Input Set : A:\Bml41018.app

Output Set: N:\CRF4\02112003\J009030A.raw

```

63 ttg ctc cat ggt tgg ccc ggc agc ttc gtt gag ttc tac cca atc ctg 384
64 Leu Leu His Gly Trp Pro Gly Ser Phe Val Glu Phe Tyr Pro Ile Leu
65      115      120      125
67 cag cta ttc cgg gag gag tac acc cct gag act ctg cca ttc cat ctg 432
68 Gln Leu Phe Arg Glu Glu Tyr Thr Pro Glu Thr Leu Pro Phe His Leu
69      130      135      140
71 gtt gtt ccg tcc ctt cct ggg tat act ttt tca tct ggt ccc ccg ctg 480
72 Val Val Pro Ser Leu Pro Gly Tyr Thr Phe Ser Ser Gly Pro Pro Leu
73 145      150      155      160
75 gac aag gac ttc ggc ttg atg gac aac gcc cgg gtc gta gac cag ttg 528
76 Asp Lys Asp Phe Gly Leu Met Asp Asn Ala Arg Val Val Asp Gln Leu
77      165      170      175
79 atg aag gac ctc ggg ttc gga agt ggt tat att att cag gga ggt gat 576
80 Met Lys Asp Leu Gly Phe Gly Ser Gly Tyr Ile Ile Gln Gly Gly Asp
81      180      185      190
83 att ggt agc ttt gtt gga cga ctg ttg ggc gtg ggt ttc gac gcc tgc 624
84 Ile Gly Ser Phe Val Gly Arg Leu Leu Gly Val Gly Phe Asp Ala Cys
85      195      200      205
87 aaa gcg gtt cat ttg aac ctg tgc gca atg agg gct ccc cct gag ggc 672
88 Lys Ala Val His Leu Asn Leu Cys Ala Met Arg Ala Pro Pro Glu Gly
89      210      215      220
91 ccg tca att gag agc ttg tcc gca gcg gag aag gag gga atc gcg cga 720
92 Pro Ser Ile Glu Ser Leu Ser Ala Ala Glu Lys Glu Gly Ile Ala Arg
93 225      230      235      240
95 atg gag aag ttc atg acc gat ggc tta gct tat gcc atg gag cac agt 768
96 Met Glu Lys Phe Met Thr Asp Gly Leu Ala Tyr Ala Met Glu His Ser
97      245      250      255
99 act ccg ccc agt act att ggc cac gtg ctg tcc agc agt ccg atc gca 816
100 Thr Arg Pro Ser Thr Ile Gly His Val Leu Ser Ser Ser Pro Ile Ala
101      260      265      270
103 tta ctt gca tgg att ggt gag aaa tat ctc caa tgg gtg gat aaa ccc 864
104 Leu Leu Ala Trp Ile Gly Glu Lys Tyr Leu Gln Trp Val Asp Lys Pro
105      275      280      285
107 ctc cct tct gag acc atc ctc gag atg gtg agc ctg tat tgg ctg acg 912
108 Leu Pro Ser Glu Thr Ile Leu Glu Met Val Ser Leu Tyr Trp Leu Thr
109      290      295      300
111 gaa agt ttc ccg cgg gca att cat acc tac cgc gag act acc cca act 960
112 Glu Ser Phe Pro Arg Ala Ile His Thr Tyr Arg Glu Thr Thr Pro Thr
113 305      310      315      320
115 gcc tcc gct ccc aat gga gcg aca atg ctt cag aag gaa tta tat att 1008
116 Ala Ser Ala Pro Asn Gly Ala Thr Met Leu Gln Lys Glu Leu Tyr Ile
117      325      330      335
119 cac aag ccg ttt ggg ttc tcc ttc ttc ccc aag gac ctt tgt cct gtg 1056
120 His Lys Pro Phe Gly Phe Ser Phe Phe Pro Lys Asp Leu Cys Pro Val
121      340      345      350
123 cct ccg agc tgg att gct aca acg gga aat cta gta ttc ttc ccg gat 1104
124 Pro Arg Ser Trp Ile Ala Thr Thr Gly Asn Leu Val Phe Phe Arg Asp
125      355      360      365
127 cat gca gag gga gga cac ttt gcc gca ttg gag cgt cca cgc gag ctg 1152

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/10/009,030A

TIME: 14:33:45

Input Set : A:\Bml41018.app

Output Set: N:\CRF4\02112003\J009030A.raw

```

128 His Ala Glu Gly Gly His Phe Ala Ala Leu Glu Arg Pro Arg Glu Leu
129      370                      375                      380
131 aag acc gac ctg aca gca ttt gtc gag cag gtg tgg cag aag tag      1197
132 Lys Thr Asp Leu Thr Ala Phe Val Glu Gln Val Trp Gln Lys
133 385                      390                      395
136 <210> SEQ ID NO: 2
137 <211> LENGTH: 398
138 <212> TYPE: PRT
139 <213> ORGANISM: Aspergillus niger
141 <400> SEQUENCE: 2
142 Met Ser Ala Pro Phe Ala Lys Phe Pro Ser Ser Ala Ser Ile Ser Pro
143   1                      5                      10                      15
145 Asn Pro Phe Thr Val Ser Ile Pro Asp Glu Gln Leu Asp Asp Leu Lys
146      20                      25                      30
148 Thr Leu Val Arg Leu Ser Lys Ile Ala Pro Pro Thr Tyr Glu Ser Leu
149      35                      40                      45
151 Gln Ala Asp Gly Arg Phe Gly Ile Thr Ser Glu Trp Leu Thr Thr Met
152      50                      55                      60
154 Arg Glu Lys Trp Leu Ser Glu Phe Asp Trp Arg Pro Phe Glu Ala Arg
155   65                      70                      75                      80
157 Leu Asn Ser Phe Pro Gln Phe Thr Thr Glu Ile Glu Gly Leu Thr Ile
158      85                      90                      95
160 His Phe Ala Ala Leu Phe Ser Glu Arg Glu Asp Ala Val Pro Ile Ala
161      100                     105                     110
163 Leu Leu His Gly Trp Pro Gly Ser Phe Val Glu Phe Tyr Pro Ile Leu
164      115                     120                     125
166 Gln Leu Phe Arg Glu Glu Tyr Thr Pro Glu Thr Leu Pro Phe His Leu
167      130                     135                     140
169 Val Val Pro Ser Leu Pro Gly Tyr Thr Phe Ser Ser Gly Pro Pro Leu
170 145                      150                      155                      160
172 Asp Lys Asp Phe Gly Leu Met Asp Asn Ala Arg Val Val Asp Gln Leu
173      165                      170                      175
175 Met Lys Asp Leu Gly Phe Gly Ser Gly Tyr Ile Ile Gln Gly Gly Asp
176      180                      185                      190
178 Ile Gly Ser Phe Val Gly Arg Leu Leu Gly Val Gly Phe Asp Ala Cys
179      195                      200                      205
181 Lys Ala Val His Leu Asn Leu Cys Ala Met Arg Ala Pro Pro Glu Gly
182      210                      215                      220
184 Pro Ser Ile Glu Ser Leu Ser Ala Ala Glu Lys Glu Gly Ile Ala Arg
185 225                      230                      235                      240
187 Met Glu Lys Phe Met Thr Asp Gly Leu Ala Tyr Ala Met Glu His Ser
188      245                      250                      255
190 Thr Arg Pro Ser Thr Ile Gly His Val Leu Ser Ser Ser Pro Ile Ala
191      260                      265                      270
193 Leu Leu Ala Trp Ile Gly Glu Lys Tyr Leu Gln Trp Val Asp Lys Pro
194      275                      280                      285
196 Leu Pro Ser Glu Thr Ile Leu Glu Met Val Ser Leu Tyr Trp Leu Thr
197      290                      295                      300
199 Glu Ser Phe Pro Arg Ala Ile His Thr Tyr Arg Glu Thr Thr Pro Thr

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/10/009,030A

TIME: 14:33:45

Input Set : A:\Bml41018.app

Output Set: N:\CRF4\02112003\J009030A.raw

```

200 305          310          315          320
202 Ala Ser Ala Pro Asn Gly Ala Thr Met Leu Gln Lys Glu Leu Tyr Ile
203          325          330          335
205 His Lys Pro Phe Gly Phe Ser Phe Phe Pro Lys Asp Leu Cys Pro Val
206          340          345          350
208 Pro Arg Ser Trp Ile Ala Thr Thr Gly Asn Leu Val Phe Phe Arg Asp
209          355          360          365
211 His Ala Glu Gly Gly His Phe Ala Ala Leu Glu Arg Pro Arg Glu Leu
212          370          375          380
214 Lys Thr Asp Leu Thr Ala Phe Val Glu Gln Val Trp Gln Lys
215 385          390          395
218 <210> SEQ ID NO: 3
219 <211> LENGTH: 21
220 <212> TYPE: DNA
221 <213> ORGANISM: Artificial Sequence
223 <220> FEATURE:
224 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
226 <400> SEQUENCE: 3
227 atgcgatcgg actgctggac a 21
230 <210> SEQ ID NO: 4
231 <211> LENGTH: 20
232 <212> TYPE: DNA
233 <213> ORGANISM: Artificial Sequence
235 <220> FEATURE:
236 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
238 <400> SEQUENCE: 4
239 cgcgggcaat ccacacctac 20
242 <210> SEQ ID NO: 5
243 <211> LENGTH: 35
244 <212> TYPE: DNA
245 <213> ORGANISM: Artificial Sequence
247 <220> FEATURE:
248 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
250 <400> SEQUENCE: 5
251 cggaattcca tggtcactgg aggagcaata attag 35
254 <210> SEQ ID NO: 6
255 <211> LENGTH: 24
256 <212> TYPE: DNA
257 <213> ORGANISM: Artificial Sequence
259 <220> FEATURE:
260 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer
262 <400> SEQUENCE: 6
263 ttgaattccc tacttctgcc acac 24
266 <210> SEQ ID NO: 7
267 <211> LENGTH: 32
268 <212> TYPE: DNA
269 <213> ORGANISM: Artificial Sequence
271 <220> FEATURE:
272 <223> OTHER INFORMATION: Description of Artificial Sequence: Primer

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/10/009,030A

TIME: 14:33:45

Input Set : A:\Bm141018.app

Output Set: N:\CRF4\02112003\J009030A.raw

274 <400> SEQUENCE: 7
275 gctgaattca catgtccgct ccgttcgcca ag 32
278 <210> SEQ ID NO: 8
279 <211> LENGTH: 44
280 <212> TYPE: DNA
281 <213> ORGANISM: Artificial Sequence
283 <220> FEATURE:
284 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
285 oligonucleotide
287 <400> SEQUENCE: 8
288 ccatgggaat tctcgagatc taagcttatg catcagctgc atgg 44

VERIFICATION SUMMARY

PATENT APPLICATION: **US/10/009,030A**

DATE: 02/11/2003

TIME: 14:33:46

Input Set : **A:\Bm141018.app**

Output Set: **N:\CRF4\02112003\J009030A.raw**

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date